中秋放送 | 数据算法基础回顾

2021-09-22 梁正霖

《数据分析思维课》 课程介绍 >

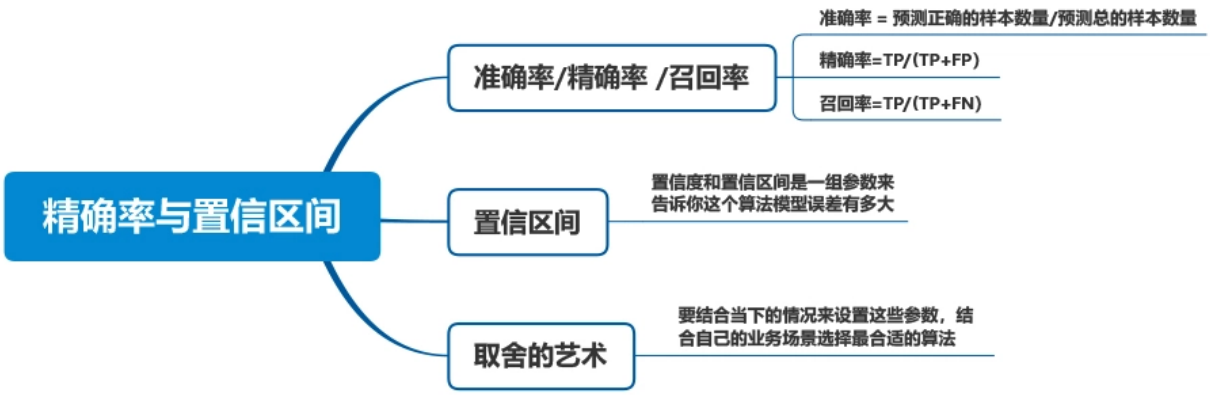
讲述：正霖
时长 00:23 大小 371.01K

你好，我是这门课程的编辑正霖。

今天是收假回来的第一天，不知道你中秋节过得怎么样呢？上节课我带你重新梳理了一遍数据分析基础部分的脉络，我们今天继续来复习回顾数据算法基础部分。另外，如果你想看到其他任何形式的加餐，欢迎在评论区留下你的需求。

话不多说，我们直接开始吧！

12 | 精确率与置信区间：两种预测，你究竟应该相信哪一个？

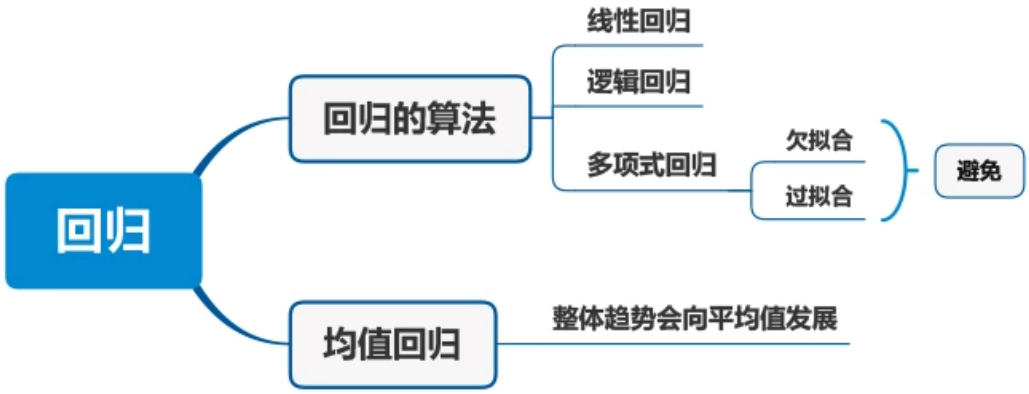


极客时间

在具体场景里，没有十全十美的算法，总是要做一些取舍。这个时候我们就更要去理解这些指标背后的含义，做到“断舍离”，根据我们实际的业务场景去选择最优的算法。

生活和工作做决策的时候也是如此，现实世界里很少有“两好选其优”的机会，大部分都是“两害取其轻”。究竟哪个害处更大不可接受，我们要自己衡量好。这样才可以在我们自己的生活和工作当中逐步的优化自己生活工作的算法，提高我们自己生活的最终的精确率，召回率和置信区间。

13 | 趋势分析与回归：父母高，孩子一定高么？

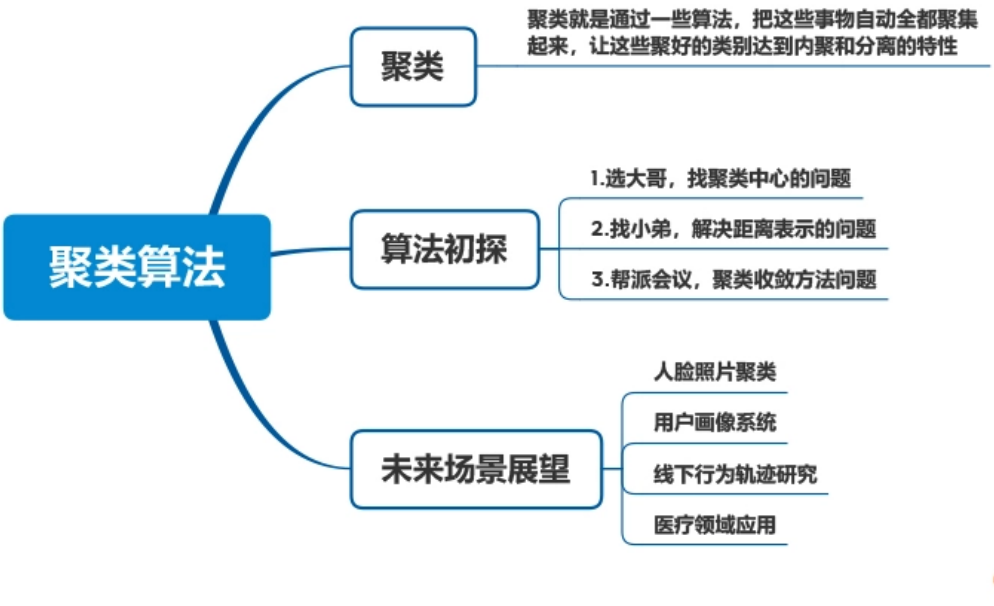


极客时间

在生活和工作里，我们可以通过回归分析找到很多简单的规律，它们能够帮助我们去预测一些常见的数据问题。但是在真正使用的时候，我们也不能盲目相信算法模型推导出来的结果，因为现实要比我们预测出来更加的贴近于平庸：好的没有我们预测当中的那么好，差的也没预测当中的那么差。

所以对我们自己的工作和生活来讲，用一颗平常心不断去提高自己的平均线水平才是正确选择。人和人之间的差异没有那么大，不存在着优生学，也不存在着“龙生龙凤生凤，老鼠的儿子会打洞”这样的说法。

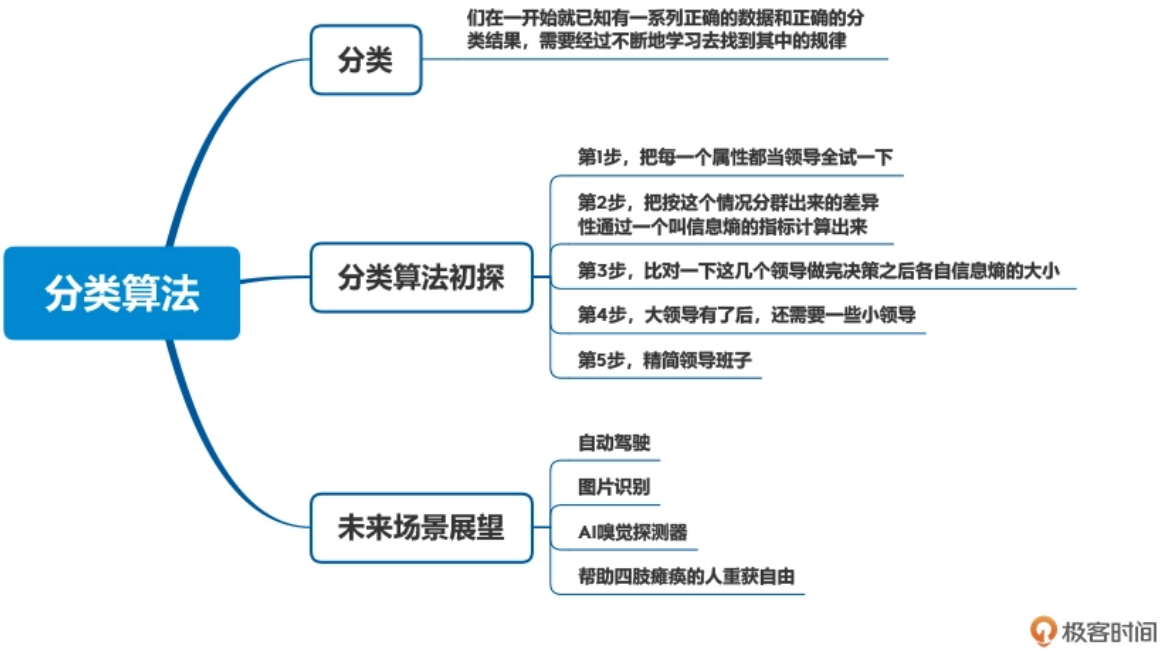
14 | 初识聚类算法：物以类聚，让复杂事物简单化



物以类聚，这是我们天生就有的一种思维方式。不知道你会不会有这样一种感觉：总是每天很忙，全都深入在不同的生活、业务细节里，但一天忙下来，却不知道自己到底忙了些什么。

你不妨试试用类似聚类算法的思路，把你觉得纷纷扰扰的一些小事统一归到一个篮子里去，用一整块的时间或者类似通用的方法去解决它们，说不定会有奇效。因为多用聚类算法的方式去思考，可以把你的思维锻炼得更加结构化，助你更快理清琐碎的生活。

15 | 初识分类算法：分而治之，不断进化

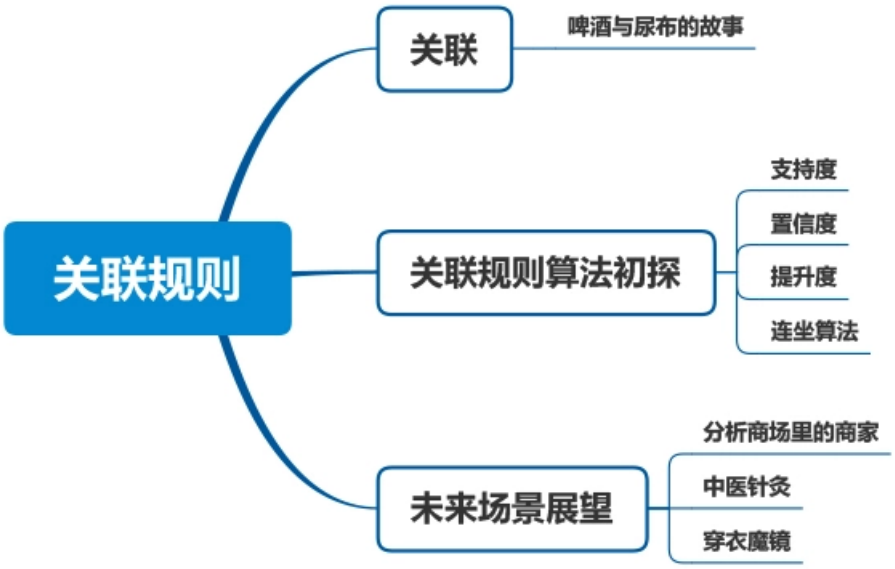


希望你能够学会用分类算法的视角去看待“复盘”这件事。

分类算法的核心就是在于经验不断积累，不断迭代自己的规则，从而得到最好的答案。而我们在工作和生活当中，其实就接触的场景和得到反馈的结果来说，要比电脑当中的分类算法多得多。但我们有像分类算法一样，把这些场景和反馈结果分类整理记录下来，然后下次遇到情况时再去优化么？其实我相信大多数的人都是没有的。

所以很多人会经历了很多事情后依然庸庸碌碌。我们需要让大脑这个超级分类器不仅去接收好结果、差结果，还要在结果之外找背后的原因来不断优化自己的算法。那些成功的人，就是通过不断地思考，不断地学习优化自己的思维，最终他们的大脑进化成为超级分类器当中的佼佼者，通过现象看到了本质。

16 | 关联规则：为什么啤酒和尿布一起卖？

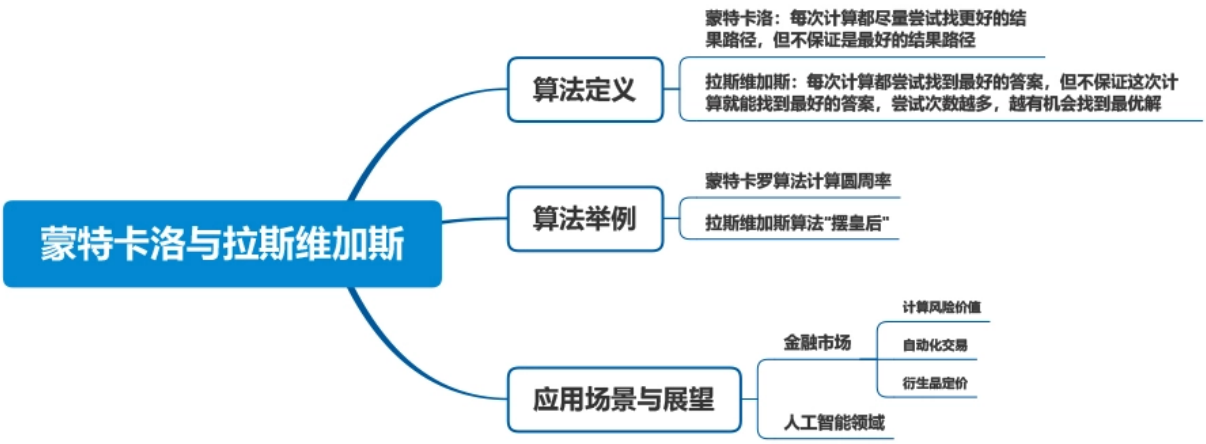


人的一生其实很短暂，我们会经历很多事情，感觉很多的事情有关联又不关联，就比如说眼皮跳真的和你今天的运气关联性很高吗？

我们要和关联算法一样，把和你关联关系最强的那些事情把握住，把关联不强的事情舍弃掉。你可以试试用关联算法的思想，盘一盘你现在手里的资源，看看能不能用“连坐”算法把整体无关的事务、人脉做到断舍离，留下精力把和你最强的关联关系的事情做好。

如果你分不清什么事情对你关联关系最强，什么事情对你无关紧要，你的生活很有可能变成一团毛线球，不知道从哪下功夫，就算发力也有可能忙活一些不痛不痒的小事。人的一生重要的事情和重要的人脉可能就这么几个，你抓住了，人生才能成功。

17 | 蒙特卡洛与拉斯维加斯：有限时间内如何获得最优解？

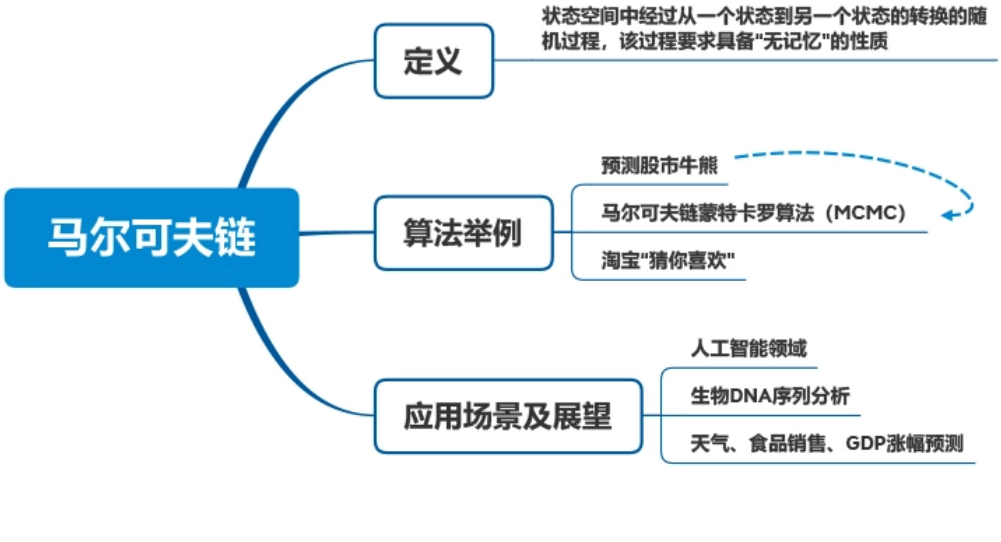


我们要根据自己手里的资源、时间，灵活选择是使用蒙特卡罗式的方法还是拉斯维加斯式的方法来处理事情，毕竟人的一生时间和精力都是有限的。

我们每天都会面临到各种各样的问题，你可以仔细思考一下你现在手里哪些事情不达目的誓不罢休（拉斯维加斯式算法）；哪些事情需要精益的方法和思路，多次尝试不断修正，事情发展到一定程度见好就收（蒙特卡洛式算法）。

对于管理企业来说，我们要高维度思考，不要把我们的有限的时间和精力浪费在不必要的事情上，整体的做事思路是抓大放小。而重要的事情要用拉斯维加斯算法一通到底，任何细节都不要放过，确保随机事件的正确性。

18 | 马尔可夫链：你的未来，只取决于你当下做什么

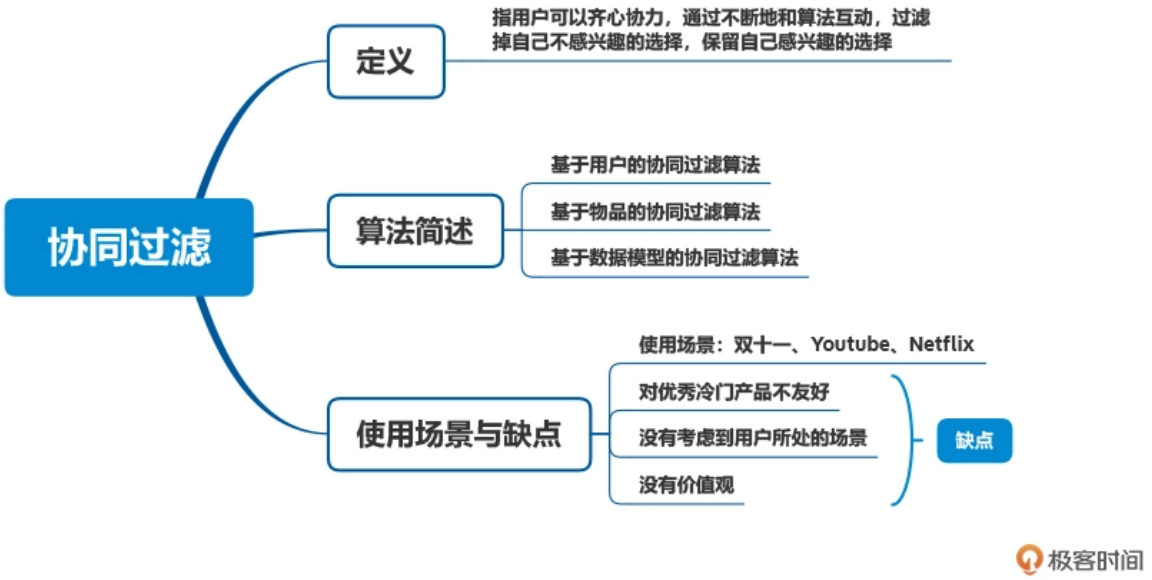


在我们工作和日常生活当中也有很多“马尔可夫链”：你现在的状态其实大部分都是由你上一个状态决定，没有人会走背字一直失败，也没有人能幸运到一直成功。

你可以仔细想想，真正的失败，很多时候都是自己遇到失败后从此一蹶不振，走不出来失败的这个状态才造成的。“没有迈不过去的坎”这句话用马尔可夫链的视角来看，那就是现在自己的状态，只和自己上一个状态相关，和整体无关。所以吸取完教训后，调整好现在的心态，用现在去影响你的未来。

我特别喜欢《飘》电影结尾郝思嘉说的那句话，我觉得它诠释了“马尔可夫链”在生活哲学中的真谛：“Tomorrow is another day”——你的未来只取决于你当下在做什么，而不是过去你曾经做过什么，毕竟“明天，是新的一天”。

19 | 协同过滤：你看到的短视频都是集体智慧的结晶

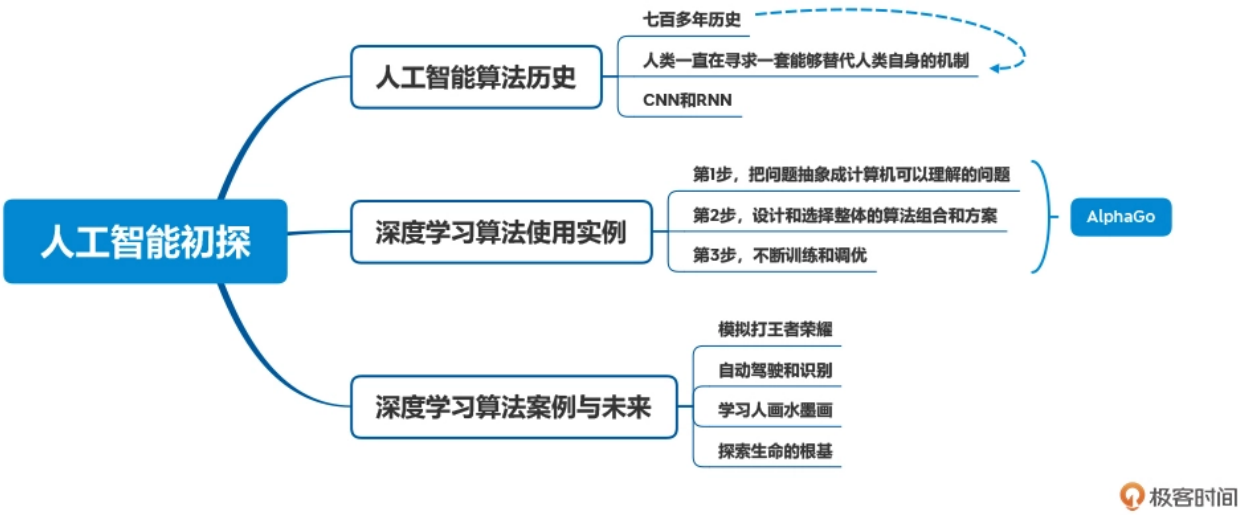


首先，你自己的心态应该更加开放，不要一股脑地追主流，毕竟主流和大众的不一定是最适合自己的，我们的圈子当中应该有个性化的东西。

同样，我们的价值观也应该更加地开放，不能就沉浸在自己的小圈子里。因为协同过滤给我们的都是我们所喜欢的东西，它的价值观并不一定是最好的，我们应该开放心态去接受和尝试各种各样新的主流的非主流的物品，用我们自己的经历和人生去判断。

我们更不要沉浸在某些短视频或者网站根据我们兴趣推荐的碎片化文章里，因为它给我们带来的不只是推荐，还会束缚、固化我们的思维，让我们成为这个时代里的“井底之蛙”。毕竟我们要主导自己的人生，而不是让算法去主导我们的人生。

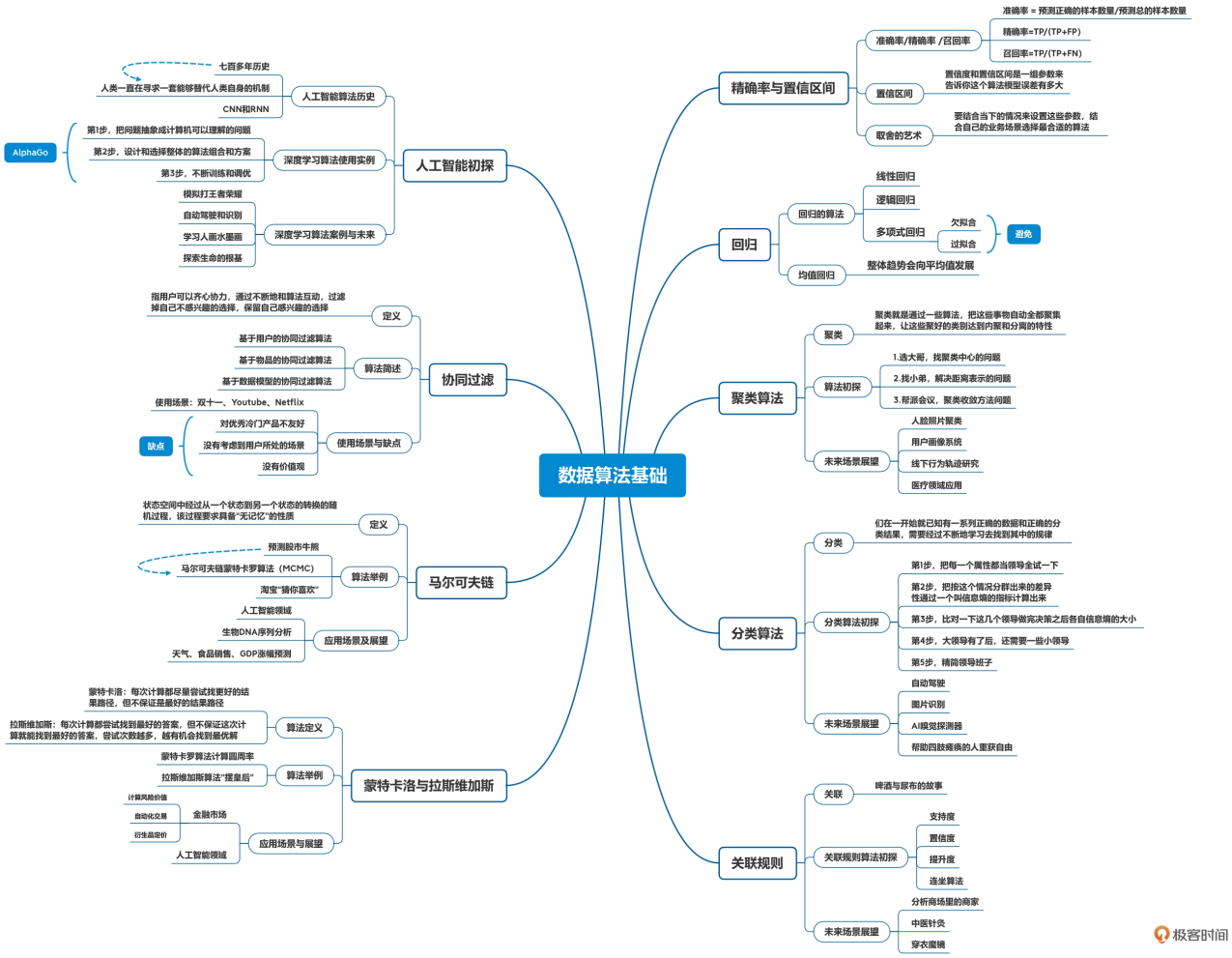
20 | 人工智能初探：阿尔法狗是怎样的一只“狗”？



尽管人工智能算法可以在很多有规则的竞争里超过人类，甚至现在很多人对人类十分悲观，觉得总有一天人工智能会像电影里那样奴役人类，但我一直都不这样认为。因为人工智能算法是没有灵魂的，它所有的计算其实本质上还是一个分类模拟器。

人工智能算法是一个有监督的学习算法，无论通过什么样的方式去模拟，它都无法通过一个有规则的算法去适配当今无规则的现实世界，更无法去模拟人们的感情、灵感和创造力。所以我们不要“机械”地活着，要往生活里多注入一些热爱和创新才好。

数据算法基础部分整合版思维导图



分享给需要的人，Ta订阅后你可得 20 元现金奖励

生成海报并分享

赞 1 提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 中秋放送 | 数据分析基础回顾

下一篇 国庆放送 | 书单推荐

精选留言 (2)

写留言

进化菌
2021-09-22

感谢整理回顾~
知识这东西啊，就是得经常翻翻，才能在需要用的时候显得没那么费力。

展开

6

小呀么小二郎
2021-09-25

等我自己能整理出来一份类似资料的时候，这门课我应该就真的学会了吧